

曖昧性を持ったユーザ発話に対する 格フレームを用いた聞き返し発話候補の生成

古川 智雅¹ 吉野 幸一郎^{1,2} 須藤 克仁¹ 中村 哲¹

¹奈良先端科学技術大学院大学情報学研究科

²科学技術振興機構 さきがけ

{furukawa.tomomasa.fo1, koichiro, sudoh, s-nakamura}@is.naist.jp

1 はじめに

近年、音声対話によって人間と情報を授受するアプリケーションが盛んに開発されている。そういったアプリケーションでは、目的によらず自然言語を理解することが必要であり、そのために近年、述語項構造解析が注目されている [1, 2]。しかし、最新の研究でも対話の音声認識結果に対する述語項構造解析の精度は60%程度と低い [3]。音声対話では話者間で共有されていることや自明なことは省略される傾向にある。そうした語の省略によって発話文上に解析上の曖昧性が含まれることが、解析精度低下の原因の一つと考えられる。対話における語の省略として、以下の例のように格と項が省略されている状態を想定する。以下の例では「何の」というノ格の項が省略されている。

(1) (Φノ) 鍋料理が食べたいです。

このような曖昧性を持ったユーザ発話に対して、テキストに対する既存の述語項構造解析では静的に解析を行う。これに対し、音声対話ではユーザに聞き返しを行うことができるため、動的に曖昧性を解消することができる。そこで本研究では、語の省略によるユーザ発話の言語理解の曖昧性を解消するために、ユーザに聞き返しを行うような対話システムの構築手法を提案する。本稿では特に、格フレーム [4] を用いて曖昧性解消のための聞き返し発話候補を生成する手法を示し、評価を行う。この際、聞き返しが必要かどうかの検知も同時に行う。

2 関連研究

音声対話においてシステムがユーザに対して聞き返しを行う手法として、述語ごとに必須格 [5] を構築し

た辞書を用いることで、不足格に対して聞き返しを行う手法 [6] がある。この研究では応答生成モジュールの一つとして相槌や繰り返などとともに聞き返しを行っており、聞き返しに対するユーザの反応から情報を得ることに主眼を置いたものではない。また、聞き返しのみに着目すると、精度が60%弱と低く、改善の余地がある。

また、ユーザとの音声対話において聞き返しの必要性検知を行う手法として、アノテーションしたデータに対して n-gram などの言語的特徴量を利用して検知を行う手法 [7] がある。この研究では聞き返しの必要性のみに焦点を当てており、具体的な聞き返しの内容を考慮していない。また、聞き返し必要性検知の精度が40%強と低く、改善の余地がある。

3 聞き返し発話候補の生成

3.1 概要

本研究が提案する対話システムでは、語の省略由来する曖昧性を持つユーザ発話に対して、述語項構造解析を行った後、格フレームを用いて聞き返し発話候補を生成する。今回は音声対話でよく起こる格と項の省略について扱う。実際のユーザ発話を考慮した場合、述語に対して省略される格と項がそれぞれ一つずつとは限らない。例えば、「鍋料理が食べたいです。」というユーザ発話に対する聞き返し発話候補として、ノ格に対する聞き返しを行う場合を考える。このとき鍋の種類について聞き返すのであれば「何の鍋料理が食べたいですか?」と聞き返し、鍋料理を作った人について聞き返すのであれば「誰の鍋料理が食べたいですか?」と聞き返す必要がある。こうしたユーザ発話に含まれる述語に対して、システムが取得したい情報に

人: 誰	組織・団体: 何	動物: 何	植物: 何	動物-部位: どこ
植物-部位: どこ	人工物-食べ物: 何	人工物-衣類: 何	人工物-乗り物: 何	人工物-金銭: どれくらい
人工物-その他: 何	自然物: 何	場所-施設: どこ	場所-施設部位: どこ	場所-自然: どこ
場所-機能: どこ	場所-その他: どこ	抽象物: 何	形・模様: 何	色: 何
数量: どれくらい	時間: いつ			

表 1: カテゴリと疑問詞を一对一で対応付けした辞書

応じて、複数の聞き返し発話候補を生成することを本研究の目標とする。例えばユーザ発話が「鍋料理が食べたいです。」の場合、聞き返し発話候補は「何の鍋料理が食べたいですか?」、「どこで食べたいですか?」などとなる。

3.2 手順

聞き返し発話候補の生成手順を以下に示す。ここでは入力のユーザ発話が「鍋料理が食べたいです。」である場合の出力例も併せて示す。

手順1 格フレームを用い意味カテゴリと格のペアを抽出
まずユーザ発話に含まれる述語に対して、KNP¹を用いた格・構文解析結果により発話文に含まれない格を抽出する。その後、これらの発話文から検出されなかった格に対して格フレームを参照する。これにより、当該述語に対し存在し得る格の項候補を得ることができる。これらの項候補に対して形態素解析器 JUMAN++² で意味カテゴリを付与すると、意味カテゴリと、格フレームの基となったコーパス上での格フレーム頻度が得られる。これに対し、頻度が多い意味カテゴリほど必要な情報であると仮定し、格の頻度と意味カテゴリの頻度の割合をしきい値として、聞き返しに用いる意味カテゴリを選択した。これにより、述語に対応する聞き返し発話候補である意味カテゴリと格のペアが抽出できる。この段階での出力結果は〈人工物, ノ〉、〈場所-施設, デ〉などとなる。ここで格の抽出の際に用いる格は、ガ格・ニ格・ヲ格・ト格・デ格・カラ格・ヨリ格・マデ格・ヘ格・ノ格の10個の格に限定する。

手順2 カテゴリを疑問詞に置き換えて疑問詞と格のペアを取得
JUMAN++で得られる、名詞の形態素に付与される意味カテゴリ 22種³と疑問詞を一对一で対

応付けした辞書を事前に作成し、これを用いてカテゴリを疑問詞に置き換えることで、疑問詞と格のペアを取得する。ここで意味カテゴリと疑問詞を一对一で対応付けした辞書を表1に示す。手順1の出力結果と表1より、この段階での出力結果は〈何, ノ〉、〈どこ, デ〉などとなる。

手順3 聞き返し発話候補を生成

手順2で取得した疑問詞と格のペアをもとに聞き返し発話候補を生成する。具体的にはノ格以外の場合は「疑問詞」+「格」+「述語の繰り返し」+「疑問形」を、ノ格の場合は「疑問詞」+「格」+「述語に係る名詞・格の繰り返し」+「述語の繰り返し」+「疑問形」をそれぞれ聞き返し発話候補とする。したがって手順2の出力結果より、出力結果は「何の鍋料理が食べたいですか?」、「どこで食べたいですか?」などとなる。

なお、「わかりました。」などの述語が含まれていても聞き返す必要がないユーザ発話も存在する。そこで、述語が一つも格と項を持たない場合は聞き返す必要がないユーザ発話であると判定し、聞き返し発話候補の生成を行わないものとした。また、格フレームの項の意味カテゴリの割合が少ない場合は聞き返す必要がないものと仮定し、それぞれの格の頻度に対して10%以上の頻度を持つ意味カテゴリに対して聞き返し発話候補を生成するものとした。例えばそれぞれガ格の項の全体の頻度が1,000、〈人工物, ガ〉のペアの頻度が150、〈場所-施設, ガ〉のペアの頻度が50であったとすると、〈人工物, ガ〉のペアの頻度はガ格全体の10%以上にあたるので聞き返し発話候補を生成するが、〈場所-施設, ガ〉については聞き返し発話候補を生成しないものとする。

4 実験

本実験では音声対話テキストコーパスとして京都観光案内対話コーパス [8] を用いた。京都観光案内対話コーパスはガイドと旅行者との音声対話をテキストに書き起こしたものである。コーパスには 63,533

¹<http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

²<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN++>

³<http://lotus.kuee.kyoto-u.ac.jp/nl-resource/jumanpp/jumanpp-manual-1.0.pdf>

聞き返し候補の有無	ユーザ発話の数
有	7,877(33.8%)
無	15,451(66.2%)
合計	23,328

表 2: 提案手法で聞き返し候補が生成されたユーザ発話

ラベル \ 予測	要	否	Recall
要	22	4	84.6%(22/26)
否	30	44	59.5%(30/74)

表 3: ランダムに選択した 100 個のユーザ発話に対する聞き返しの必要性検知の評価結果

個の発話が含まれ、そのうち述語を含む発話は全体の 36.7%にあたる 23,328 個であった。これは音声対話コーパスの中に、述語を含まない相槌や繰り返しのユーザ発話が多かったためと考えられる。本実験では、その 23,328 個の発話に対して聞き返し発話候補の生成実験を行った。本実験の統計情報を表 2 に示す。23,328 個の発話のうち、提案手法が聞き返し発話候補を生成したユーザ発話の数は 33.8%にあたる 7,877 個であった。

本実験では、聞き返しの必要性検知の精度を表すカバー率 (Recall) と聞き返し発話候補自体の正しさ (Precision) の二つの評価指標で評価を行った。Recall はランダムに選択した 100 個のユーザ発話に対して、聞き返しが必要な場面で質問が生成できたかについて主観評価した。Precision は実際に生成された聞き返し発話候補に対し、候補として正しい聞き返し発話候補を生成できているかについて主観評価した。

4.1 聞き返しの必要性検知の評価 (Recall)

ユーザ発話に対して主観的に聞き返しが必要と判断した 26 発話、聞き返しが必要ないと判断した 74 発話をテストセットとして評価を行った。その結果聞き返しの必要性検知の精度は 66%となり、先行研究 [7] よりも高い Recall で聞き返し発話を生成することができた。以下に正しく検知できたユーザ発話の例を示す。

- (2) a. 天龍寺前もありますけれど、バス停がですね、いくつかありますので、十一系統のバスに乗っていただきまして、
b. 電車とバスね。はい、わかりました。

それぞれ (2-a) は聞き返し発話候補を生成したユーザ発話の例、(2-b) は聞き返し発話候補を生成しなかつ

たユーザ発話の例である。(2-a) ではユーザ発話に対して「どこまで乗ればいいですか?」といった聞き返しが必要となるため、聞き返しが必要という判定をした。一方、(2-b) はユーザ発話が繰り返してであるために聞き返す必要がないというラベルを付与した。これに対し提案手法は、ユーザ発話に述語を含んでいるが、述語が一つも格と項を含んでいないために聞き返し発話候補を生成しなかった。

次に正しく検知できなかったユーザ発話の例を以下に示す。

- (3) a. 小さな滝が三つ並んでるんですけどね。そこで、
b. 三年坂、二年坂歩かれて。
c. ただ、わりと京都市バスが発達してますのでね、

(3-a)、(3-b) は聞き返し発話候補を生成する必要があるにもかかわらず生成しなかった例、(3-c) は生成する必要がないにもかかわらず生成した例である。(3-a) では意味カテゴリーの頻度がしきい値よりも少なかったため、(3-b) では、述語「歩かれて」の格と項が一つもないと解析されたためにそれぞれ聞き返し発話候補を生成することができなかった。これらの例に対して聞き返し発話候補を生成させるためには、頻度が少なくても格フレームの項の中に疑問詞が含まれている場合は聞き返すといった、聞き返し発話候補の生成条件を緩和することが必要であると考えられる。また、(3-b) については述語と格フレームが出現形のまま扱われているため頻度が少なくなったと考えられ、原形への正規化によって問題を軽減することができると考えられる。一方 (3-c) ではユーザ発話のみでも十分な情報が含まれているにもかかわらず、「何に発達してますのでか?」といった聞き返し発話候補を生成した。こうした判断には格や述語の情報だけではなく対話の文脈を考慮する必要がある。

4.2 聞き返し発話候補の評価 (Precision)

4.1 で正しく聞き返す必要があると判断できたユーザ発話 22 個に対して、システムが生成した聞き返し発話候補 81 個を評価した。生成された質問文が対話の文脈上必要な質問文か主観で正誤評価した。生成した聞き返し発話候補の Precision は 32.1%(26 個/81 個)となり、先行研究より低くなった。以下に一つのユーザ発話に対して正しく生成できた例と誤って生成された例を示す。なお、今回は生成された文の不自然さは

評価対象外とする。これは生成に用いたルールがかなり単純であるためであり、生成された文の不自然さの改善は今後の課題である。

- (4) a. 最初バス降りたのは、この五条坂からですね。ずっと歩いて、こうお寺まで行ったんですけどもね。
- (i) 誰と行ったんですか？
- (ii) 何に行ったんですか？
- (iii) 何で行ったんですか？

それぞれ (4-a-i) は正しい聞き返し発話候補の生成例、(4-a-ii) と (4-a-iii) は誤った聞き返し発話候補の生成例である。(4-a-i) ではト格の情報を得るために聞き返し発話候補を生成しており適切である。一方 (4-a-ii) ではユーザがお寺に行ったことは自明であるため、(4-a-iii) ではユーザが歩いてお寺に行ったことは自明であるためそれぞれ誤っていると判定した。(4-a-ii) や (4-a-iii) では「行った」という述語に対する格の情報としては欠落しているが、他の述語やそれに対する格などの他の情報から補完することができる。このような他の情報から補完可能な場合に対して、誤って聞き返し発話候補を生成した例が多く見られた。こういったユーザ発話に対しては、パラフレーズすることで明示的に項と格を用いてユーザ発話を書き直すといった工夫が必要であると考えられる。また、聞き返しに用いるかどうかを対話制御の機構で制御するという手法も考えられる。

また、聞き返し発話候補の生成方法が異なるノ格に対して、生成した聞き返し発話候補の正しい例と誤った例をそれぞれ示す。

- (5) a. の店もありますし、
- (i) 何の店がありますか？
- (6) a. そこから階段で下りられると、音羽の滝ってというのがありまして、
- (i) 何ののがありますか？

(5-a-i) ではユーザ発話の情報が欠落しており、正しく聞き返し発話候補を生成している。一方 (6-a-i) ではノ格の項の情報を得るために聞き返し発話候補を生成する段階までは正しいが、ノ格が名詞「の」に係るのは不自然であるため誤りと判定した。このように、ノ格では係り先の名詞による誤りが見られるため、ノ格の係り先の名詞が格フレームに含まれるかどうかを生成基準として新たに用いることを検討している。

5 おわりに

本稿では格フレームを用いて曖昧性を持つユーザ発話に対して聞き返し発話候補を生成する手法を提案した。実験ではユーザ発話に対して 66% の Recall で、曖昧性を持つユーザ発話に対して聞き返し発話候補の生成を行うことができた。また、生成発話の内容については 32.1% の Precision で正しい聞き返し発話候補を生成した。今後の課題として精度の向上やコンテキストを考慮した聞き返し発話候補の生成、実際に対話システムに用いる場合に利用するかどうかの制御を行うことが挙げられる。

謝辞

本研究の一部は JST CREST(課題番号: JP-MJCR1513) および JST さきがけ (課題番号: JP-MJPR165B) の支援を受けて行った。

参考文献

- [1] Koichiro Yoshino, Shinsuke Mori, and Tatsuya Kawahara. Spoken dialogue system based on information extraction using similarity of predicate argument structures. *Annual SIGdial Meeting on Discourse and Dialogue*, pp. 59–66, 2011.
- [2] Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chikaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino, and Yoshihiro Matsuo. Towards an open-domain conversational system fully based on natural language processing. In *COLING*, pp. 928–939, 2014.
- [3] 今村賢治, 東中竜一郎, 泉朋子. 対話解析のためのゼロ代名詞照応解析付き述語項構造解析. *自然言語処理*, Vol. 22, No. 1, pp. 3–26, 2015.
- [4] 河原大輔, 黒橋禎夫. 高性能計算環境を用いた web からの大規模格フレーム構築. *情報処理学会研究報告自然言語処理 (NL)*, Vol. 2006, No. 1, pp. 67–73, 2006.
- [5] C. J. Fillmore. *The case for case*. 1967.
- [6] 下岡和也, 徳久良子, 吉村貴克, 星野博之, 渡部生聖. 音声対話ロボットのための傾聴システムの開発. *自然言語処理*, Vol. 24, No. 1, pp. 3–47, 2017.
- [7] 大原康平, 佐藤翔悦, 吉永直樹, 豊田正史, 喜連川優. 不足情報を自律的に問う対話エージェントの実現に向けた聞き返しの必要性検討. 第9回データ工学と情報マネジメントに関するフォーラム (DEIM フォーラム 2017), 2017.
- [8] 翠輝久, 大竹清敬, 堀智織, 柏岡秀紀, 中村哲. 京都観光案内対話コーパスにおける対話行為タグの設計と分析. *情報処理学会研究報告音声言語情報処理 (SLP)*, Vol. 2009, No. 10, pp. 39–44, jan 2009.